

# Application of principal component analysis in grouping geomorphic parameters of a watershed for hydrological modeling

S. K. Sharma · S. Gajbhiye · S. Tignath

Received: 6 October 2013 / Accepted: 23 February 2014 / Published online: 20 March 2014  
© The Author(s) 2014. This article is published with open access at Springerlink.com

**Abstract** Principal component analysis has been applied to 13 dimensionless geomorphic parameters on 8 sub-watersheds of Kanhiya Nala watershed tributary of Tons River located in Part of Panna and Satna district of Madhya Pradesh, India, to group the parameters under different components based on significant correlations. Results of principal component analysis of 13 geomorphic parameters clearly reveal that some of these parameters are strongly correlated with the components but texture ratio and hypsometric integral do not show correlation with any of the component. So they have been screened out of analysis. The principal component loading matrix obtained using correlation matrix of eleven parameters reveals that first three components together account for 93.71 % of the total explained variance. Therefore, principal component loading is applied to get better correlation and clearly group the parameters in physically significant components. Based on the properties of the geomorphic parameters, three principal components were defined as drainage, slope or steepness and shape components. One parameter each from the significant components may form a set of independent parameters at a time in modeling the hydrologic responses such as runoff and sediment yield from small watersheds.

**Keywords** Geomorphic parameters · Principal component analysis · GIS

## Introduction

Watershed is an ideal unit for planning and management of land and water resources (Gajbhiye et al. 2013). It is a natural hydrological entity which allows surface runoff to a defined channel, drain, stream or river at a particular point (Chopra et al. 2005). Physiography, drainage, geomorphology, soil, land use/land cover are some of the parameters which play a significant role in watershed planning (Javed et al. 2011). Watershed management involves proper utilization of land, water, forest and soil resources. Therefore, realistic assessment of the hydrological behavior of a watershed is important to develop effective management plan. There may be various considerations for the implementation of management programs in the few sub-watersheds only. It is always better to start management measures from the most critical sub-watershed. Sediment yield from a catchment is one of the main criteria to find most critical sub-watershed to soil erosion. However, this criterion requires for assessing continuous monitoring of sediment samples at the catchment outlet. Such data are hardly available in India for small watersheds. Although the sediment yield from large basins can be obtained from such observation, it is not possible to ascertain the vulnerability to soil erosion of small watersheds within a basin. In the absence of sediment yield data morphometric parameters may be helpful in assessing most critical sub-watershed.

Morphometry is the measurement and mathematical analysis of the configuration of the earth's surface, shape and dimensions of its landform (Clarke 1966). This analysis can be achieved through measurement of linear, aerial

S. K. Sharma (✉)  
Department of Soil and Water Engineering, College of  
Agricultural Engineering, J.N.K.V.V., Jabalpur (M.P.), India  
e-mail: sharmashailesh501@gmail.com

S. Gajbhiye  
Department of Water Resources Development and Management,  
Indian Institute of Technology, Roorkee (U.K.), India

S. Tignath  
Department of Geology, Government Science College, Jabalpur  
(M.P.), India

and relief aspects of basin and slope contributions. Morphometric analysis of a basin can be better achieved through a latest technology like RS (Remote Sensing) and Geographical Information System (GIS) as conventional measurement of these parameters is laborious and cumbersome. Many researchers have demonstrated the potential of RS and GIS technique for morphometric analysis of watershed (Shrimali et al. 2001; Thakker and Dhiman 2007; Sharma et al. 2010).

The method of quantitative analysis of watershed was developed by Horton (1945) and was further modified by Strahler (1964). Sufficient works on the quantitative analysis of geomorphological parameters of watersheds have been done in India and abroad (Ghose et al. 1969). However, a very little work on the interrelationship of morphological parameters has been carried out. To determine interrelationship of these geomorphological parameters is very important to develop sediment yield regression models (Hydrological modeling). Statistical methods are applied in a variety of fields in hydrological research. Factor analysis is useful for interpreting morphometric parameters and relating the same to specific hydrological processes. Multivariate analysis is simply a collection of procedures for analyzing the associations between two or more sets of data that have been collected on each object in one or more samples of object. Synder (1962) introduced some solutions, possibilities of multivariate statistics in hydrological modeling. Wong (1979) utilized a multivariate statistical technique component analysis in analyzing the effects of twelve basins and climatological parameters. Wallis (1965) in discussion of multivariate statistical methods in hydrology recommends, for multifactor hydrological problems, the use of principle component analysis with varimax rotation of the factor weight matrix. Haan and Allen (1972), Decoursy and Deal (1974) have also demonstrated the use of multiple regression analysis for development of hydrological prediction equations involving geomorphic parameters. Mishra and Satyanarayana (1988) carried out principal component analysis with varimax rotation on ten geomorphic parameters at Damodar Valley catchment of India and concluded that nine parameters could be significantly grouped into three components. Singh et al. (2009) carried out principal component analysis to 13 geomorphic parameters collected for sixteen watersheds of Chambel catchment of Rajasthan. The parameters are grouped into three components. Therefore, in this study an attempt has been made to determine geomorphological parameters and to study the intercorrelation (multicollinearity) among variables to screen out the less significant variables out of the analysis and to arrange the remaining into physically significant groups by applying principal component analysis for better interpretability.

## Materials and methods

The study area Kanhiya Nala watershed lies within the Tons River catchment  $80^{\circ}32'24''$  to  $80^{\circ}34'12''$ E longitude and  $24^{\circ}06'$  to  $24^{\circ}10'48''$ N latitude (Fig. 1) with elevation ranges from 480 to 600 m above mean sea level and extends a total area of 25.58 km<sup>2</sup>. The average annual rainfall is 1,300 mm. The watershed is situated in Satna and Panna district of Madhya Pradesh.

The Survey of India (SOI) toposheet number 63 D/12 on 1:50,000 scale was used to prepare a base map. The satellite data of IRS-P6, LISS-III sensor with 23.5 m resolution was used in the present study to prepare updated drainage map. False Color Composite (FCC) of study area is presented in Fig. 2.

For generation of digital input maps, image processing and digital analysis of data, Arc GIS 9.1 and ERDAS Imagine 9.1 software are used in the present study. SPSS 14.0 is also used for statistical analysis.

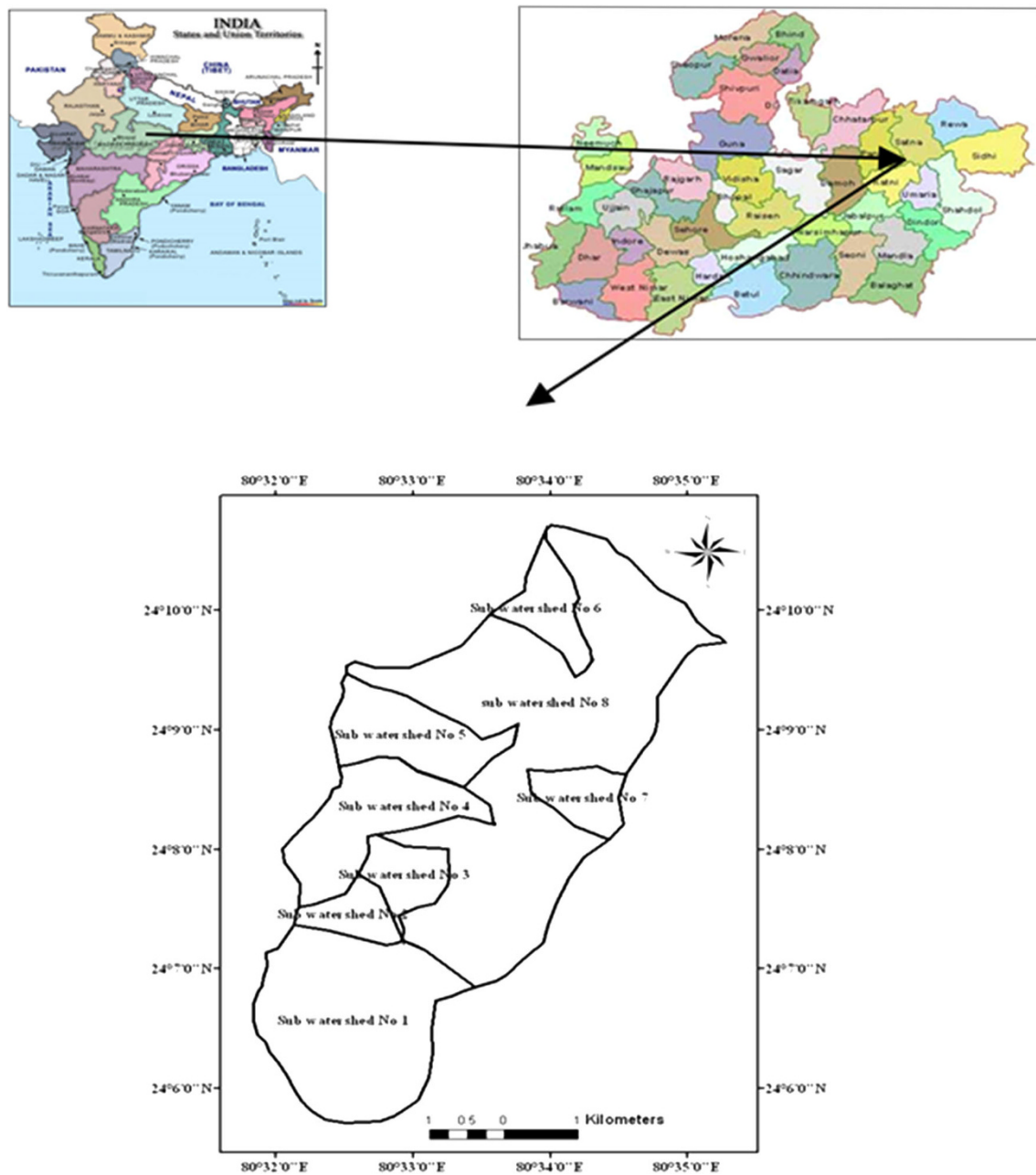
### Watershed delineation from the topological data

The topological information of the study area was digitized from georeferenced SOI toposheets 63 D/12 using the ArcGIS tools. The watershed boundary, sub-watershed boundary, stream network and contours were digitized in vector form to generate digital input maps. The drainage was initially digitized from SOI toposheets and later updated using IRS-P6 LISS-III Satellite data. The digitized sub-watershed boundary, updated stream network and contour lines were used for further geomorphological analysis (Figs. 3, 4).

### Geomorphic parameters

Geomorphologic characteristics play a vital role on the hydrological response from a watershed, and therefore, a number of parameters which signify the watershed characteristics are evaluated from toposheets in GIS environment. For the present study, entire catchment of the Kanhiya Nala which is a tributary of Tons river of Madhya Pradesh, India was discretized into eight sub-watersheds. The input parameters for present study such as area, perimeter, stream order, number of streams, stream length, elevation and basin length were derived from digitized stream network and contour map in GIS environment. The geomorphic parameters for the discretized watershed area were calculated using formula suggested by Horton (1945), Strahler (1964), Schumm (1956) and Miller (1953) given in Table 1.

Hypsometric analysis of drainage basin is carried out to develop the relationship between horizontal cross-sectional drainage basin area and the elevation. In analysis, a curve is derived by plotting the relative height ( $h/H$ ) and relative



**Fig. 1** Location map of study area

areas ( $a/A$ ); the obtained curve is called as hypsometric curve (Suresh 1997). The shape of hypsometric curve varies in early geologic stages of development of the drainage basin, but once a steady state is attained it tends to vary little despite lowering relief (Kumar 1991; Suresh 1997).

#### Principal component analysis

The method of principal components or component analysis is based upon the early work of Pearson with specific adaption to principal component analysis suggested by the

Hotelling (1933). The geomorphometric parameters are usually many times correlated. The correlation indicates that some of the information contained in one variable is also contained in some of the other remaining variables. More specifically, the first principal component is that linear combination of the original variables which contributes a maximum to their total variance; the second principal component, uncorrelated with the first, contributes a maximum to the residual variance, and so on until the total variance is analyzed. Since the method is so dependent on the total variance of the original variables, it

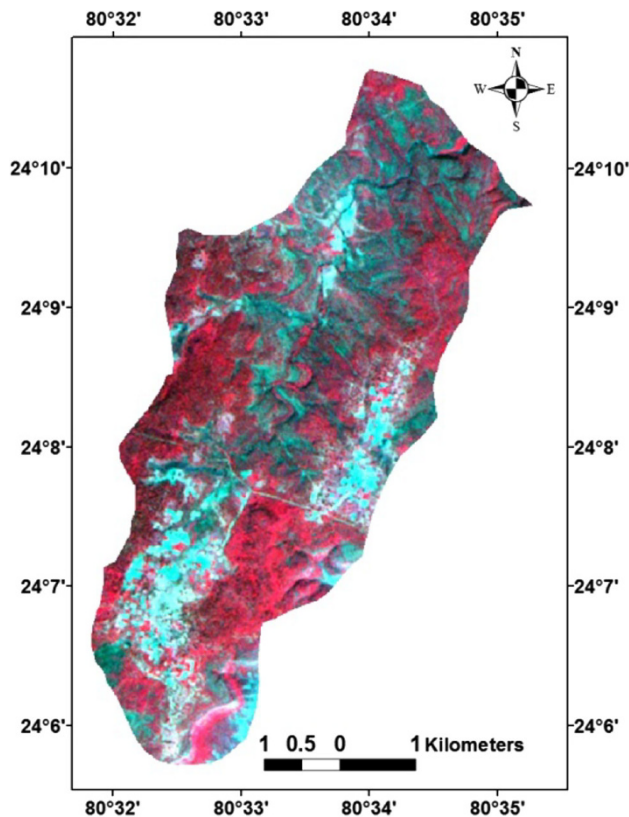


Fig. 2 False Color Composite of Study area

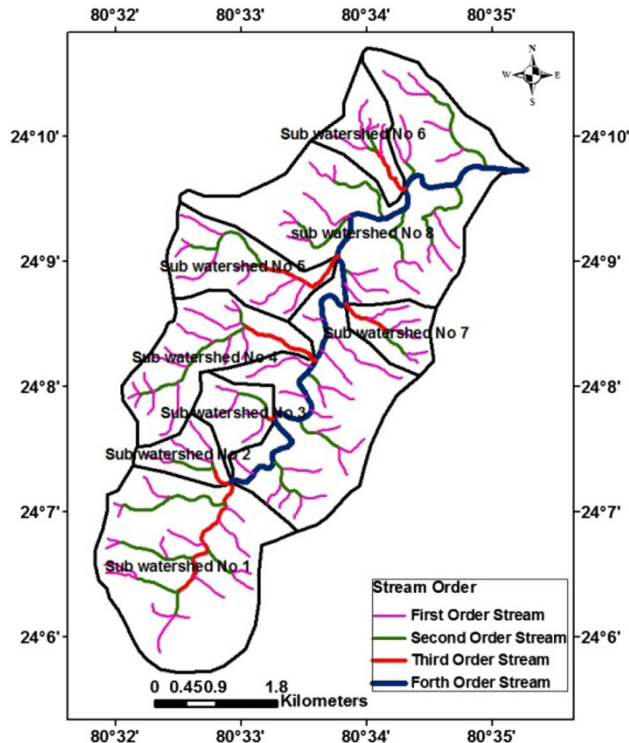


Fig. 3 Stream network of study area

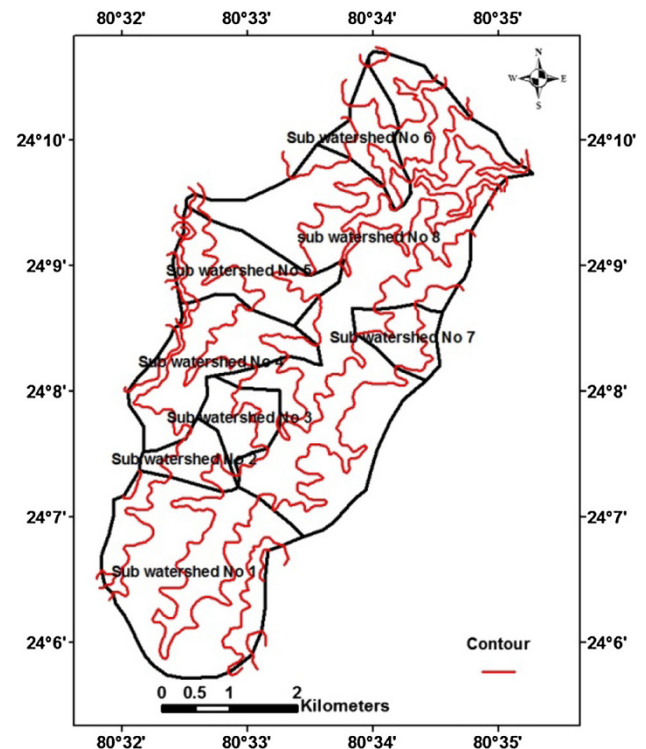


Fig. 4 Contour map of study area

is most suitable when all the variables are measured in the same units. Hence, it is customary to express the variables in standard form, i.e., to select the unit of measurement for each variables so that its sample variance is one. Then, the analysis is made on the correlation matrix, with the total variance equal to  $n$ . The objectives are achieved in three steps:

Step 1 Calculate the correlation matrix,  $R$ .

Step 2 Calculate the principal component loading matrix by principal component analysis.

Step 3 In the principal component (PC) Loading matrix, eigen values greater than 1 indicates significant PC loading.

Eigen value indicated how well each of the identified factors fit the data from all the geomorphic parameters on all the principal components.

#### 1. Correlation matrix

The inter-correlation matrix of the geomorphic parameters is obtained using the following procedure:

(a) The parameters are standardized

$$X = (x_{ij} - x_j) / S_j \quad (1)$$

where  $X$  denotes the matrix of standardized parameters,  $x_{ij}$   $i$ th observation on  $j$ th parameters,  $i = 1 \dots N$  (Number of observation),  $j = 1 \dots P$  (Number of observation),  $x_j$  mean of



**Table 1** Formula for computation of geomorphic parameters

| Geomorphic parameters             | Formula   | Reference     |
|-----------------------------------|---|---------------|
| Bifurcation ratio ( $R_b$ )       | $R_b = N_u/N_{u+1}$<br>where, $R_b$ = Bifurcation ratio<br>$N_u$ = Total number of stream of segment of order $u$<br>$N_{u+1}$ = Total number of stream of segment of next higher order | Schumm (1956) |
| Drainage density ( $D_d$ )        | $D_d = L_u/A$<br>where, $D_d$ = Drainage density<br>$L_u$ = Total stream length of order $u$<br>$A$ = Area of watershed   | Horton (1945) |
| Texture ratio ( $T$ )             | $T = N_1/P$<br>where, $T$ = Texture ratio<br>$N_1$ = Total number of streams of first order<br>$P$ = Perimeter  | Horton (1945) |
| Stream frequency ( $F_u$ )        | $F_u = N_u/A$<br>where, $N_u$ = Total number of streams of all order<br>$A$ = Area of watershed   | Horton (1945) |
| Circulatory ratio ( $R_c$ )       | $R_c = 4\pi A/P^2$<br>where, $R_c$ = Circulatory ratio<br>$A$ = Area of watershed<br>$P$ = Perimeter  | Miller (1953) |
| Form factor ( $R_f$ )             | $R_f = A/L_b^2$<br>where, $R_f$ = Form factor<br>$A$ = Area of watershed<br>$L_b$ = Length of watershed   | Horton (1945) |
| Elongation ratio ( $R_e$ )        | $R_e = (2/L_b) * (A/\pi)^{0.5}$<br>where, $R_e$ = Form factor<br>$A$ = Area of basin<br>$L_b$ = Length of basin   | Schumm (1956) |
| Length of overland flow ( $L_o$ ) | $L_o = \frac{1}{2} D_d$<br>where, $L_o$ = Length of overland flow<br>$D_d$ = Drainage density   |               |
| Relative relief ( $R_r$ )         | $R_r = H/P$<br>where, $R_r$ = Relative relief<br>$H$ = Maximum watershed relief<br>$P$ = Perimeter of basin   |               |
| Relief ratio ( $R_h$ )            | $R_h = H/L_b$<br>where, $R_h$ = Relief ratio<br>$H$ = Maximum watershed relief<br>$L_b$ = Length of basin   |               |
| Ruggedness number ( $R_N$ )       | $R_N = H * D_d$<br>where, $R_N$ = Ruggedness number<br>$H$ = Maximum watershed relief<br>$D_d$ = Drainage density   |               |
| Compactness Coefficient           | $C_c = 0.2821 * A/P^{0.5}$<br>where, $C_c$ = Compactness Coefficient<br>$A$ = Area of basin<br>$D_d$ = Drainage density   |               |

the  $j$ th parameters,  $S_j$  Standard deviation of the  $j$ th parameters.

- (b) The correlation matrix of parameters is the minor product moment of the standardized predictor measures divided by  $N$  and is given by

$$R = (x' \times x)/N \quad (2)$$

where,  $x'$  denotes the transpose of the standardized matrix of predictor parameters

## 2. Principal component loading matrix

The principal component loading matrix which reflects how much a particular parameter is correlated with different factors, is obtained by premultiplying the characteristics vector with square root of the characteristics values of the correlation matrix.

Thus,

$$A = Q \times D^{0.5} \quad (3)$$

where  $A$  principal component loading matrix,  $Q$  characteristics vector of the correlation matrix,  $D$  characteristics value of the correlation matrix.

## Result and discussion

Morphometric parameters of sub-watersheds were calculated in GIS environment and are presented in Table 2 and computed geomorphometric parameters are presented in Table 3.

The correlation matrix (Table 4) of 13 geomorphic parameters reveals that strong correlations (correlation coefficient more than 0.9) exist between bifurcation ratio, form factor and elongation ratio, between drainage density and length of overland flow, between circulatory ratio and compactness coefficient, between form factor and elongation ratio and between relative relief and relief ratio. Also, good correlation (correlation coefficient more than 0.75) exists between bifurcation ratio, stream frequency and ruggedness number, between texture ratio relief ratio and hypsometric integral, between stream frequency form factor and elongation ratio, between circulatory ratio and ruggedness number and between circulatory ratio and ruggedness number. Some more moderately correlated parameters exist (correlation coefficient more than 0.60) between bifurcation ratio ruggedness number, between drainage density and stream frequency, between texture ratio form factor and elongation ratio, between stream frequency and length of overland flow, between form factor and circulatory ratio, between circulatory ratio and elongation ratio, between relative relief and hypsometric integral and between relief ratio and hypsometric integral.

**Table 2** Sub-watershed wise input geomorphic parameters

| Sub-watershed No. | Area (km <sup>2</sup> ) | Perimeter (km) | Length of basin (km) | Total No. of stream | Total stream length (km) | Max. elevation (m) | Min. elevation (m) |
|-------------------|-------------------------|----------------|----------------------|---------------------|--------------------------|--------------------|--------------------|
| 1                 | 5.54                    | 9.35           | 3.47                 | 22                  | 13.26                    | 600                | 560                |
| 2                 | 0.77                    | 3.84           | 1.13                 | 7                   | 3.01                     | 600                | 560                |
| 3                 | 0.93                    | 4.62           | 1.26                 | 7                   | 2.47                     | 580                | 540                |
| 4                 | 2.61                    | 8.10           | 2.26                 | 17                  | 9.09                     | 600                | 540                |
| 5                 | 1.99                    | 6.85           | 1.94                 | 10                  | 6.59                     | 600                | 520                |
| 6                 | 1.09                    | 5.53           | 1.38                 | 7                   | 2.91                     | 580                | 500                |
| 7                 | 0.89                    | 3.92           | 1.23                 | 8                   | 3.24                     | 560                | 540                |
| 8                 | 11.96                   | 27.38          | 5.37                 | 48                  | 34.33                    | 600                | 480                |

**Table 3** Sub-watershed wise computed geomorphic parameters

| Sub-watershed No. | $R_b$ | $D_d$ | $T$  | $F_u$ | $R_c$ | $R_f$ | $R_e$ | $L_o$ | $R_r$ | $R_h$ | $R_N$ | $C_c$ | $H_{si}$ |
|-------------------|-------|-------|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| 1                 | 4.13  | 2.39  | 1.82 | 3.97  | 0.80  | 0.46  | 0.77  | 0.209 | 0.004 | 0.012 | 0.096 | 1.121 | 0.39     |
| 2                 | 2.00  | 3.93  | 1.04 | 9.14  | 0.65  | 0.60  | 0.88  | 0.127 | 0.010 | 0.035 | 0.157 | 1.237 | 0.52     |
| 3                 | 2.00  | 2.64  | 0.87 | 7.49  | 0.55  | 0.59  | 0.86  | 0.189 | 0.009 | 0.032 | 0.106 | 1.347 | 0.59     |
| 4                 | 4.50  | 3.48  | 1.73 | 6.51  | 0.50  | 0.51  | 0.81  | 0.144 | 0.007 | 0.027 | 0.209 | 1.413 | 0.54     |
| 5                 | 4.50  | 3.32  | 1.17 | 5.03  | 0.53  | 0.53  | 0.82  | 0.151 | 0.012 | 0.041 | 0.265 | 1.372 | 0.52     |
| 6                 | 2.00  | 2.68  | 0.72 | 6.44  | 0.45  | 0.57  | 0.86  | 0.187 | 0.014 | 0.058 | 0.214 | 1.496 | 0.68     |
| 7                 | 2.25  | 3.62  | 1.28 | 8.96  | 0.73  | 0.59  | 0.87  | 0.138 | 0.005 | 0.016 | 0.072 | 1.169 | 0.59     |
| 8                 | 7.13  | 2.87  | 1.31 | 4.01  | 0.20  | 0.41  | 0.83  | 0.174 | 0.004 | 0.022 | 0.345 | 2.234 | 0.52     |

**Table 4** Inter-correlation matrix of 13 geomorphic parameters

| Parameters | $R_b$  | $D_d$  | $T$    | $S_f$  | $R_c$  | $R_f$  | $C_c$  | $R_e$  | $R_r$  | $R_h$  | $R_N$  | $L_o$  | $H_{si}$ |
|------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|----------|
| $R_b$      | 1.000  | −0.179 | 0.531  | −0.782 | −0.590 | −0.931 | 0.726  | −0.932 | −0.484 | −0.351 | 0.734  | 0.125  | −0.472   |
| $D_d$      | −0.179 | 1.000  | −0.009 | 0.659  | 0.150  | 0.440  | −0.200 | 0.438  | 0.103  | 0.011  | −0.013 | −0.992 | 0.107    |
| $T$        | 0.531  | −0.009 | 1.000  | −0.423 | 0.267  | −0.628 | −0.103 | −0.618 | −0.717 | −0.763 | −0.072 | 0.039  | −0.772   |
| $S_f$      | −0.782 | 0.659  | −0.423 | 1.000  | 0.381  | 0.888  | −0.476 | 0.882  | 0.241  | 0.142  | −0.543 | −0.626 | 0.485    |
| $R_c$      | −0.590 | 0.150  | 0.267  | 0.381  | 1.000  | 0.430  | −0.944 | 0.438  | −0.132 | −0.350 | −0.883 | −0.054 | −0.345   |
| $R_f$      | −0.931 | 0.440  | −0.628 | 0.888  | 0.430  | 1.000  | −0.609 | 0.988  | 0.567  | 0.442  | −0.553 | −0.413 | 0.588    |
| $C_c$      | 0.726  | −0.200 | −0.103 | −0.476 | −0.944 | −0.609 | 1.000  | −0.621 | −0.141 | 0.089  | 0.846  | 0.121  | 0.136    |
| $R_e$      | −0.932 | 0.438  | −0.618 | 0.882  | 0.438  | 0.988  | −0.621 | 1.000  | 0.573  | 0.445  | −0.555 | −0.412 | 0.584    |
| $R_r$      | −0.484 | 0.103  | −0.717 | 0.241  | −0.132 | 0.567  | −0.141 | 0.573  | 1.000  | 0.969  | 0.180  | −0.125 | 0.600    |
| $R_h$      | −0.351 | 0.011  | −0.763 | 0.142  | −0.350 | 0.442  | 0.089  | 0.445  | 0.969  | 1.000  | 0.344  | −0.052 | 0.684    |
| $R_N$      | 0.734  | −0.013 | −0.072 | −0.543 | −0.883 | −0.553 | 0.846  | −0.555 | 0.180  | 0.344  | 1.000  | −0.070 | 0.069    |
| $L_o$      | 0.125  | −0.992 | 0.039  | −0.626 | −0.054 | −0.413 | 0.121  | −0.412 | −0.125 | −0.052 | −0.070 | 1.000  | −0.170   |
| $H_{si}$   | −0.472 | 0.107  | −0.772 | 0.485  | −0.345 | 0.588  | 0.136  | 0.584  | 0.600  | 0.684  | 0.069  | −0.170 | 1.000    |

It is very difficult at this stage to group the parameters into components and attach any physical significance. Hence, in the next, the principal component analysis has been applied. The correlation matrix is subjected to the principal component analysis.

The principal component loading matrix obtained from correlation matrix (Table 5) reveals that the first three components whose eigen values are greater than 1, together

account for about 91.458 % of the total explained variance. The first component is strongly correlated (loading of more than 0.90) with circulatory ratio and compactness coefficient and moderately correlated (loading of more than 0.60) with form factor and elongation ratio, which may be termed as shape component. The second component is strongly correlated with relief ratio and ruggedness number and good correlated (loading of more than 0.80) which may

**Table 5** Principal component loading matrix of 13 geomorphic parameters

| Parameters                               | Component |        |        |
|--|-----------|--------|--------|
|  | 1         | 2      | 3      |
| $R_b$                                    | −0.117    | −0.540 | −0.825 |
| $D_d$                                    | 0.082     | 0.005  | 0.987  |
| $T$                                      | −0.029    | −0.524 | −0.019 |
| $S_f$                                    | 0.577     | 0.369  | 0.656  |
| $R_c$                                    | 0.920     | −0.296 | 0.044  |
| $R_f$                                    | 0.665     | 0.527  | 0.396  |
| $C_c$                                    | −0.935    | 0.051  | −0.088 |
| $R_e$                                    | 0.670     | 0.525  | 0.392  |
| $R_r$                                    | 0.036     | 0.883  | 0.037  |
| $R_h$                                    | −0.158    | 0.924  | −0.025 |
| $R_N$                                    | 0.136     | −0.968 | 0.038  |
| $L_o$                                    | 0.003     | −0.045 | −0.991 |
| $H_{si}$                                 | −0.028    | 0.556  | 0.160  |
| Eigen value                              | 6.186     | 3.641  | 2.062  |
| % of Total Factor Co variance            | 47.583    | 28.010 | 15.864 |
| Cumulative % of Total Factor Co variance | 47.583    | 75.594 | 91.458 |

be termed as slope or steepness component. Third component is strongly correlated with drainage density and length of overland flow, good correlation with bifurcation ratio and moderately correlated with stream frequency. It is evident from these results that some of the parameters are highly correlated with some of the components but the parameters texture ratio and hypsometric integral could not be grouped with any of the component because of their poor correlation with them.

To screen out parameters having less significance in explaining the component variance, parameters texture ratio and hypsometric integral are screened out from analysis. Then correlation matrix and principal component matrix are obtained for eleven parameters.

The principal component loading matrix obtained using the correlation matrix of eleven parameters (Table 6) reveals that the first three components now together accounts for 94.491 % of the total explained variance showing an increase of about 3.033 %.

The principal component loading here also improved considerably in almost all significant parameters. The circulatory ratio and compactness coefficient have strong correlation (loadings of more than 0.90) with the first component. The elongation ratio and form factor have moderate correlation (loadings of more than 0.60) with first component. The relative relief, relief ratio and ruggedness number have strong correlation with the second component. The bifurcation ratio, drainage density and length of overland flow have strong correlation (loadings of more

**Table 6** Principal component loading matrix of eleven finally screened out geomorphic parameters

| Parameters                               | Component |        |        |
|--|-----------|--------|--------|
|  | 1         | 2      | 3      |
| $R_b$                                    | −0.148    | −0.532 | 0.816  |
| $D_d$                                    | −0.061    | 0.009  | 0.983  |
| $S_f$                                    | −0.527    | 0.299  | 0.684  |
| $R_c$                                    | −0.914    | −0.227 | 0.042  |
| $R_f$                                    | −0.651    | 0.504  | 0.429  |
| $C_c$                                    | 0.920     | −0.040 | −0.088 |
| $R_e$                                    | −0.656    | 0.506  | 0.424  |
| $R_r$                                    | −0.002    | 0.964  | 0.053  |
| $R_h$                                    | 0.186     | 0.973  | −0.007 |
| $R_N$                                    | 0.151     | 0.978  | 0.024  |
| $L_o$                                    | −0.023    | −0.039 | −0.988 |
| Eigen value                              | 5.704     | 2.724  | 1.966  |
| % of Total Factor Co variance            | 51.852    | 24.767 | 17.872 |
| Cumulative % of Total Factor Co variance | 51.852    | 76.619 | 94.491 |

than 0.90) with third component. The stream frequency has moderate correlation (loadings of more than 0.60) with third component.

It is observed that the first component is strongly correlated with circulatory ratio and compactness coefficient and good correlation with ruggedness number which is grouped under shape component. The second component has strong correlation with relative relief, relief ratio and ruggedness number and termed as slope or steepness component. The third component has strong correlation with bifurcation ratio, drainage density and length of overland flow and moderate correlation with stream frequency hence is called as drainage component.

It can be seen how useful the principal component analysis has been in screening out the parameters or variables of least significance and regrouping the remaining variables into the physically significant factors. Multiple regression technique can then be applied in modeling the hydrological responses such as surface runoff and sediment yields from the watersheds. One parameter each from the significant components may form a set of independent parameters at a time in modeling the said hydrologic responses.

## Conclusion

In the present study, 13 geomorphic parameters were evaluated for eight discretized sub-watersheds of Kanhiya watershed located in part of Panna and Satna district of Madhya Pradesh, India for principal component analysis.

The correlation matrix of the 13 geomorphic parameters revealed that strong correlations (correlation coefficient more than 0.9) exist between bifurcation ratio, form factor and elongation ratio, between drainage density and length of over land flow, between circulatory ratio and compactness coefficient, between form factor and elongation ratio and between relative relief and relief ratio. The principal component loading matrix obtained from correlation matrix reveals that first three components whose eigen values are greater than 1, together accounts for about 91.458 of the total explained variance. Based on the results of the principal component analysis, first component is strongly correlated with circulatory ratio and compactness coefficient. The second component is strongly correlated with relief ratio and ruggedness number. However, third component is strongly correlated with drainage density and length of overland flow. The texture ratio and hypsometric integral could not be grouped with any of the component because of their poor correlation with them. After screening out these parameters, the principal component loading matrix of eleven parameters indicates that first three components together account for 94.491 % of the total explained variance. Based on the properties of the geomorphic parameters, three principal components were defined as drainage, slope or steepness and shape components. One parameter each from the significant components may form a set of independent parameters at a time in modeling the hydrologic responses such as runoff and sediment yield from small watersheds. The principal component analysis is a good tool for screening out the insignificant parameters from the analysis.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Chopra R, Dhiman RD, Sharma PK (2005) Morphometric analysis of sub-watersheds in Gurudaspur district, Punjab using remote sensing and GIS techniques. *J Indian Soc Remote Sens* 33(4):531–539
- Clarke JI (1966) Morphometry from maps. *Essays in geomorphology*. Elsevier Publishing Co, New York, pp 235–274
- Decoursy D, Deal RB (1974) General aspect of multivariate analysis with application to some problems in hydrology. In: *Proceedings of symposium on statistical hydrology*, USDA, miscellaneous publication No. 1275. Washington DC, pp 47–68
- Gajbhiye S, Mishra SK, Pandey A (2013) Prioritizing erosion-prone area through morphometric analysis: an RS and GIS perspective. *Appl Water Sci* (Springer). doi:10.1007/s13201-013-0129-7
- Ghose B, Pandey S, Singh S (1969) Quantitative geomorphology of the drainage basin in semi arid environment. *Ann Arid Zone* 1:37–44
- Haan CT, Allen DM (1972) Comparison of multiple regression and principal component regression for predicting water yields in Kentucky. *Water Resour Res* 8(6):1593–1596
- Horton RE (1945) Erosional development of streams and their drainage basins: a hydrophysical approach to quantitative morphology. *Geol Soc Am Bull* 56:275–370
- Hotelling H (1933) Analysis of complex of statistical variables into principal component. *J Educ Psychol* 24:417o–441o 498–520
- Javed A, Khamday AY, Rais S (2011) Watershed prioritization using morphometric and land use/land cover parameters: a remote sensing and GIS based approach. *J Geo Soc India* 78:63–75
- Kumar V (1991) Hydrologic response models for prediction of runoff and sediment yield from small watersheds. Unpublished Ph.D Thesis. Indian Institute of Technology, Kharagpur, India, p 350
- Miller VC (1953) A quantitative geomorphic study of drainage basin characteristics in the Clinch mountain area, Virginia and Tennessees. Department of Navy, Office of Naval Research, Technical Report 3, Project NR 389-042, Washington DC
- Mishra N, Satyanarayana T (1988) Parameter grouping—a prelude to hydrologic modeling. *Indian J Power River Val Dev* 256–260
- Schumm SA (1956) Evaluation of drainage system and slopes in bed lands at Perth Ambry, New Jersey. *Geol Soc Am Bull* 67:597–646
- Sharma SK, Rajput GS, Tignath S, Pandey RP (2010) Morphometric analysis of a watershed using GIS. *J Indian Water Res Soc* 30(2):33–39
- Shrimali SS, Aggarwal SP, Samra JS (2001) Prioritizing erosion-prone areas in hills using remote sensing and GIS—a case study of the Sukhna Lake catchment, Northern India. *Int J Appl Earth Obs and Geoinf* 2(1):54–60
- Singh PK, Kumar V, Purohit RC, Kothari M, Dashora PK (2009) Application of principal component analysis in grouping geomorphic parameters for hydrologic modeling. *Water Resour Manage* 23:325–339
- Strahler AN (1964) Quantitative geomorphology of drainage basins and channel networks. Section 4-II. In: Chow VT (ed) *Handbook of applied hydrology*. McGraw-Hill, USA, pp 4–39
- Suresh R (1997) Soil and water conservation engineering. Standard Publishers Distributors, New Delhi, p 973
- Synder WM (1962) Some possibilities for multivariate analysis in hydrologic studies. *J Geophys Res* 62(2):721–729
- Thakker AK, Dhiman SP (2007) Morphometric analysis and prioritization of miniwatersheds in Mohr watershed, Gujarat using Remote sensing and GIS techniques. *J Indian Soc Remote Sens* 35(4):313–321
- Wallis RJ (1965) Multivariate statistical methods in hydrology—a comparison using data of known functional relationship. *Water Resour Res* 1:447–467
- Wong ST (1979) A multivariate statistical model for predicting mean annual flood in New England. *Ann Assoc Am Geogr* 53:293–311